36-201 — Computer Lab #8 — Partial Solutions

Question #1: 55% of 20 games is 11 games, so I'd expect the team to win 11 games.

Question #2: For SEASON1 I got the following win/loss record:
`0 1 1 1 0 0 1 0 1 1 1 0 0 0 1 0 1 1 0 1` . This is exactly 11 wins (honest, I didn't cheat!), for a winning proportion of 11/20=0.55. Note: your answers should be different (these are random events!), but broadly similar to my answers for this and the next three questions.

♣ **Question #3:** For SEASON2 I got the following win/loss record:
`0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 1 1 0` . This is 6 wins, for a winning proportion of 6/20=0.33. The team did not have a winning season; the longest winning streak was three in a row; the longest losing streak was six in a row.

Question #4: For SEASON3, I got `1 1 1 1 1 1 1 1 1 1 1 0 0 0 1 0 0 1 1 0` . 14 wins; winning proportion is 14/20 = 0.70.

Question #5: For SEASON4, I got `0 1 1 1 1 1 0 0 1 1 0 1 0 0 0 1 1 1 1 1` . 13 wins; winning proportion is 13/20 = 0.65.

Question #6: The proportions were 0.55, 0.33, 0.70 and 0.65. They are not all the same and they are not all equal to the parameter 0.55. They are bouncing around because of *random variability.* Since each sample is random we get slightly different win/loss records from each sample, and this leads to different winning proportions. But on average they should all be close to 0.55 (at least, they will be if the samples are not biased!).

♣ **Question #7:** The average is $(0.55 + 0.33 + 0.70 + 0.65)/4 = 0.5575$, which is pretty close to the population parameter, 0.55.

This is not required for the question, but it is interesting to compute SE_p and the margin of error for proportion of games won here. $SE_p \approx \sqrt{(0.5)(0.5)/20} = 0.11$, so the margin of error is about 0.22. The intervals we get from each of the four seasons are:

$$\begin{array}{lll} \text{SEASON1:} & \text{from } 0.55 - 0.22 = 0.33, & \text{to } 0.55 + 0.22 = 0.77 \\ \text{SEASON2:} & \text{from } 0.33 - 0.22 = 0.11, & \text{to } 0.33 + 0.22 = 0.55 \\ \text{SEASON3:} & \text{from } 0.70 - 0.33 = 0.37, & \text{to } 0.70 + 0.22 = 0.92 \\ \text{SEASON4:} & \text{from } 0.65 - 0.33 = 0.32, & \text{to } 0.65 + 0.22 = 0.87 \end{array}$$

and in each season the population parameter 0.55 was indeed trapped by that season's interval. (This should actually happen in 95 out of every 100 seasons, according to the middle part of the 68–95–99.7 rule).

♣ **Question #8:** The calculation of SE_p and margin of error for an SRS of size 1031 would be:

$SE_p \approx \underline{\sqrt{0.5 \cdot 0.5/1031} = 0.0156}$
$\pm 2 \cdot SE_p \approx \underline{\pm 2 \cdot 0.0156 = 0.0311}$, or about 3%.

Question #9: This agrees with USA Today's reported margin of error of $\pm 3\%$, even though USA Today probably took a more complex sample than an SRS.

Question #10: The true p whose vote will not be affected is likely to be between $\underline{51\% - 3\% = 48\%}$ and 52% + 3% = 54%.

♣ **Question #11:** The truth is likely to be somewhere between 48% and 54%. Some percentages greater than 50% (a majority) are in this interval, but also some lower than 50% are in this interval. This means we can't say for sure whether a majority of voters will let the impeachment votes affect their vote for their own US Senator in the next election.